Congestion-aware Bi-modal Delivery Systems Utilizing Drones

Mark Beliaev[1], Negar Mehr[2], and Ramtin Pedarsani[3]

Abstract—Bi-modal delivery systems are a promising solution to the challenges posed by the increasing demand of e-commerce. Due to the potential benefit drones can have on logistics networks such as delivery systems, some countries have taken steps towards integrating drones into their airspace. In this paper we aim to quantify this potential by developing a mathematical model for a Bi-modal delivery system composed of trucks and drones. We propose an optimization formulation that can be efficiently solved in order to design socially-optimal routing and allocation policies. We incorporate both societal cost in terms of road congestion and parcel delivery latency in our formulation. Our model is able to quantify the effect drones have on mitigating road congestion, and can solve for the path routing needed to minimize the chosen objective. To accurately capture the effect of stopping trucks on road latency, we model it using SUMO by simulating roads shared by trucks and cars. Based on this, we show that the proposed framework is computationally feasible to scale due to its reliance on convex quadratic optimization techniques.

I. INTRODUCTION

The growing demand of e-commerce is causing a precipitous increase in delivery truck traffic, impacting travel time for both public and private vehicles within densely populated urban areas [1]. For the past decade, e-commerce sales have grown at a yearly rate of 15%, with an astonishing 40% increase in 2020 alone [2]. It has been shown that trucks negatively impact road congestion by changing the behaviour of passenger car drivers around them [3]. Some have looked at ways to address these impacts, finding that the best strategy is restricting all heavy vehicles from using the road [4]. In addition to increased road congestion, the upsurge in truck traffic has led to other negative consequences such as air pollution and parking space shortages [5], [6]. To address these issues, researchers have been developing Bi-modal delivery systems that incorporate parcel-delivering drones [7], [8].

Due to the aerial reach of Unmanned Aerial Vehicles, or drones, they possess an inherent versatility that makes them promising for logistics networks [9]. Europe has already taken critical steps towards integrating drones into their current airspace infrastructure [10]. In their effort to overcome the challenges posed to drone delivery services, Parcel at Deutsche Post DHL Group has demonstrated success in mul-

Fig. 1: A depiction of a Bi-modal delivery system utilizing both trucks and drones. Delivery trucks making stops along roads with less lanes are shown to have a greater effect on road congestion. On the other hand, drones are shown landing directly at their assigned delivery destinations, hence not contributing negative effects on road congestion.

tiple field tests, using their recent Parcelcopter 4.0 to deliver medicine for the 400,000 residents of the Ukerewe island district of Lake Victoria [11]. These examples show the feasibility of implementing drones within logistics networks such as parcel delivery systems, depicted in Fig. 1.

In this paper we aim to quantify the potential drones have in alleviating traffic congestion induced by parcel delivery services by developing a mathematical model for a Bi-modal delivery system composed of trucks and drones. We set out to propose an optimization formulation that can be efficiently solved in order to design socially-optimal routing and allocation policies. Our goal is to incorporate both societal cost in terms of road congestion and parcel delivery latency in our optimization formulation.

We propose a new network model consisting of two separate graphs which represent the road network and the aerial network. A requisite for our model is the ability to capture the average time it takes parcels to be delivered. In addition, we want our model to capture the societal impact that different routing and allocation policies have on the road network. This way, given an objective function posed as a trade-off between societal cost and parcel latency, we form an optimization problem to solve for the socially-optimal routing strategy.

To model the effect of stopping delivery trucks on an individual link within the road network we use the open source traffic simulation package "SUMO" [12]. Based on the simulation results, we can model the latency of a link as linear with respect to the flow of stopping and non-stopping vehicles. This way we can develop a scalable optimization

[1]Mark Beliaev is with the Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA `mbeliaev@ucsb.edu`

[2]Negar Z. Mehr is with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA `negar@illinois.edu`

[3]Ramtin Pedarsani is with the Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA `ramtin@ucsb.edu`

framework by making our objective function quadratic and convex.

The main contributions of the paper are as follows:

- We propose a new Bi-modal delivery network model consisting of delivery trucks, drones, and regular vehicles.
- We describe a scalable convex optimization framework which solves for the socially-optimal allocation and routing scheme for our network model.
- We study the effect stopping trucks have on roads using SUMO simulations, and derive a mathematical latency function for roads shared between regular vehicles and delivery trucks.
- We perform numerical evaluations on our results to show the effect drone delivery systems have on road congestion and parcel delivery time.

II. RELATED WORK

In the following, we discuss a few areas of work related to ours.

A. Drone Delivery Systems

The application of drones within delivery systems has received notable interest [13]. Lee [14] has looked at the limitations posed to drones by their battery capacity and limited payload. Other work has looked at the facility location problem, optimizing for the locations of drone refueling stations in a large urban area [15]. More recently, a case study of Amazon prime air in the city of San Francisco has demonstrated how a proposed heuristic algorithm can approximately solve this NP-hard facility location problem [16]. Although these works consider the management and design of drone delivery systems as a whole, the optimizations performed do not consider the actual routing of drones within the network and their impact on the traffic network.

B. Vehicle Routing Problem

Solving for optimal routing schemes of a parcel delivery system within a road network can be viewed as an example of a vehicle routing problem (VRP) [17], [18]. These are typically solved with mixed integer linear programming (MILP) formulations which scale poorly in real world applications. On top of this, Dorling et al. [19] have shown that naively applying existing VRPs to systems utilizing drones can lead to inefficient results. One notable work formulates the hybrid vehicle-drone routing problem for pick-up and delivery services as a MILP [20], but they do not consider the affect the proposed routes have on society, and they do not allow trucks to dispatch and deliver packages. Although there are other works that consider relevant variants of the VRP [21], [22], the goal in these problems is to minimize the time it takes to complete the required routes, with no concern for the impact these schedules have on society.

C. Multi-modal Route Planning

There has been considerable progress within the past decade in developing new techniques for route planning in transportation networks [23], [24], [25], [26]. Specifically, we have seen work that incorporates drones in the traffic model

and optimizes for their routes, while allowing drones to utilize multiple modes of transportation [27], [7], [28], [29]. Other works have considered using drones in tandem with delivery trucks, without exploiting public transit [30], [31]. Although related to our problem, none of these works consider the effect drones have on the transportation network from a societal perspective. By decoupling the routes drones can take within the aerial network from the routes delivery trucks can take within the road network, we focus on the impact drones can have on road congestion as well as parcel latency within the road network.

III. NETWORK MODEL

In our work, we model the delivery modes as two separate networks, without interlinked paths between them. Each network is modeled by a directed graph composed of the same set of nodes $\mathcal{N}$ and two different sets of edges $\mathcal{E}^R$ and $\mathcal{E}^A$, where $G^R = (\mathcal{N}, \mathcal{E}^R)$ and $G^A = (\mathcal{N}, \mathcal{E}^A)$ represent the road and aerial networks respectively. We assume that each edge $e \in \mathcal{E}^R, \mathcal{E}^A$ connects two distinct nodes (v, w), representing a directed edge from v towards w. Additionally, we assume that both graphs are connected and denote node v_0 as the main delivery hub, making all other nodes possible delivery destinations. Finally, we restrict each edge $e \in \mathcal{E}^A$ to only using v_0 as the origin node, since the versatility of drones should allow them to have direct links between the main delivery hub and the destinations. We depict these two graphs in Fig. 2 for clarity.

To simplify notation, we introduce the sets of paths $\mathcal{P}^R$ and $\mathcal{P}^A$, where $p \in \mathcal{P}^R$ denotes a path within G^R and $p \in \mathcal{P}^A$ denotes an aerial path within G^A. For the road network, each path $p \in \mathcal{P}^R$ connects the delivery hub v_0 to one of the remaining nodes using a collection of edges $e \in \mathcal{E}^R$. For the aerial network, each path $p \in \mathcal{P}^A$ only contains one edge $e \in \mathcal{E}^A$ since G^A is a star network. Furthermore, for each
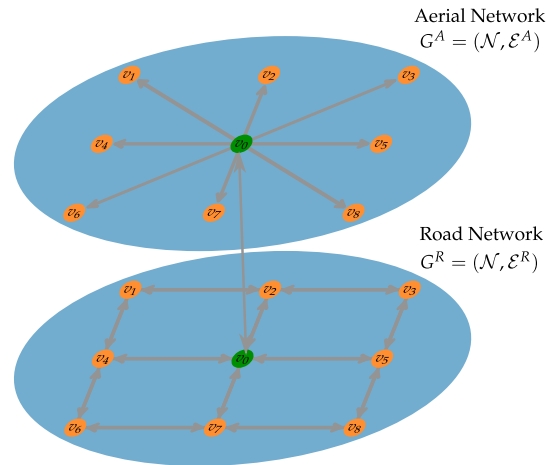

Fig. 2: This figure visualizes the proposed network model. We see that the general structure of the directed graph G^R allows us to model typical transportation networks. On the other hand, the graph G^A is given the topology of a star graph since we allow drones to take direct paths to their destinations within the aerial network.

node $v \in \mathcal{N}$, we define the subsets $\mathcal{U}_v^R, \mathcal{D}_v^R \subset \mathcal{E}^R$ which correspond to the set of edges arriving at v and the set of edges departing from v, respectively.

A. Demand

In our model, we let d_v represent the parcel demand of node $v \in \mathcal{N}$, measured in parcels per hour. Note that we actually use the set $\mathcal{N} \backslash \{v_0\}$ since the delivery hub does not have any demand, but we leave this distinction to the reader in order to avoid cumbersome notation. To ensure package delivery at all destinations, for each node $v \in \mathcal{N}$, we must have:

$$d_v = d_v^R + d_v^A, \tag{1}$$

where d_v^R and d_v^A correspond to the components of node v's parcel demand satisfied by trucks within the road network and drones within the aerial network, respectively. How these components are derived is explained in the following section.

B. Flow

In contrast to the model's delivery demands measured in parcels per hour, we express the individual path and edge flows in units of vehicles per hour. For the road network, we use f_p^T to represent the flow of parcel-carrying trucks along a given path $p \in \mathcal{P}^R$. Moreover, we let f^T be the vector of truck flows along all the paths $p \in \mathcal{P}^R$. For edge $e \in \mathcal{E}^R$, we use $f_e^T = \sum_{p \in \mathcal{P}^R : e \in p} f_p^T$ to denote the flow of trucks along edge e. Additionally, each edge $e \in \mathcal{E}^R$ has a corresponding flow of non-delivering vehicles f_e^0 which are constants defined by the model. We refer to f_e^0 as the *nominal* vehicular flow along edge e to outline that it is not a decision variable.

For the road network, we assume that each truck carries m parcels. Thus, at each node v, the number of parcels delivered by trucks is:

$$\frac{1}{m} d_v^R = \sum_{e \in \mathcal{U}_v^R} f_e^T - \sum_{e \in \mathcal{D}_v^R} f_e^T. \tag{2}$$

We can see from (2) that d_v^R is measured as the *decrease* in truck flow passing through node v, since the difference between incoming and outgoing delivery trucks determines the number of parcels dropped off at node v. Note that we do not consider the effect returning trucks have on the traffic network as they no longer need to make stops and are typically a negligible portion of the overall flow on any given edge.

For the aerial network, we let f_p^D denote the flow of drones along path $p \in \mathcal{P}^A$. Due to the limited payload of drones [32], we assume that each drone delivers only one package. This assumption coupled with the aforementioned restriction which allows only one aerial path to each destination node, makes flow f_p^D equivalent in magnitude to parcel demand d_v^A when p is the aerial path that arrives at destination node v. We note that f^T is used as our decision variable, deriving the value for f_p^D from (1) and (2) as the leftover parcel demand that needs to be delivered through drones.

C. Parcel Latency

We would like to derive the mean latency experienced by parcels in the system. We express this average latency in units of hours as:

$$L = \frac{L^R + L^A}{\sum_{v \in \mathcal{N}} d_v}, \tag{3}$$

where we let L^R and L^A represent the sums of experienced latencies for parcels delivered within the road network and aerial network respectively. This way, dividing by the total demand in (3) gives us the average latency we want. We go on to describe how these two components of latency are derived.

To derive the latency component L^R, we first need to compute the individual edge latencies ℓ_e^R for all edges $e \in \mathcal{E}^R$. Once these are computed, we can express the sum of latencies experienced by parcels delivered via trucks as:

$$L^R = \sum_{p \in \mathcal{P}^R} m f_p^T \left(\sum_{e \in p} \ell_e^R \right), \tag{4}$$

where the term in the parenthesis is the latency along a path p. As will be shown later, the edge latencies ℓ_e^R are a linear function of the two flow components on that edge, making the expression for L^R quadratic with respect to our control variable f^T. Details of how this function was calculated are saved for Section V describing our SUMO simulations. Note that (4) is only valid when all trucks are carrying the same number of parcels m. As long as the demand of individual vertices is large enough, the majority of trucks will be full, making this inaccuracy negligible.

For the drone network, we note that drone flows f_p^D do not have an effect on the latencies of the aerial paths since in the regimes of interest, they will cause no aerial congestion. Thus, we model all drones traveling along aerial path $p \in \mathcal{P}^A$ to have a latency ℓ_p^A equal to a predefined constant. This gives us the following expression for the sum of latencies experienced by parcels delivered via drones:

$$L^A = \sum_{p \in \mathcal{P}^A} f_p^D \ell_p^A, \tag{5}$$

where we no longer need to sum over the edges in a given path as done in (4) since each aerial path consists of only one edge.

It is important to note that we are not considering the possibility of failed deliveries among the road and aerial networks. This extra consideration can be implemented into our formulation using a probabilistic approach, and defining a constant μ_e to denote the probability of failure along a given link $e \in \mathcal{E}^R$ within the road network. The same can be done for the aerial paths $p \in \mathcal{P}^A$.

D. Societal Cost

Due to their recurrent stops, utilizing a large amount of delivery trucks can negatively impact traffic congestion as a whole, especially when considering package deliveries in dense urban environments [1], [3], [4], [5], [6]. To capture this effect, we define the societal cost L^S, as the normalized latency experienced by the nominal flow of traffic in units of hours:

$$L^S = \frac{1}{\beta} \sum_{e \in \mathcal{E}^R} f_e^0 \ell_e^R, \tag{6}$$

where β represents the total flow of nominal vehicles within the network in units of vehicles per hour. Since the edge latencies ℓ_e inside the summation of (6) are weighted by the corresponding nominal flows f_e^0, we are giving higher weight to roads that have a larger amount of affected drivers. We normalize this sum by the total nominal flow β so that our societal cost L^S represents a normalized latency per driver rather than an aggregate sum. We use the expressions "societal cost" and "societal latency" interchangeably when referring to L^S.

E. Operational Cost

Finally, we model the cost of operating this delivery system, which we denote with C in units of dollars per hour. We break up the cost into two components associated with the road and aerial network. To calculate these components, we define c^T and c^D as the hourly cost of operating one truck and drone respectively. Because $\frac{1}{m}d_v^R$ and d_v^A represent node v's total hourly demand of trucks and drones, we can multiply these vehicle demands with their respective hourly rates to derive the operational cost:

$$C = \frac{c^T}{m}\sum_{v \in \mathcal{N}} d_v^R + c^D \sum_{v \in \mathcal{N}} d_v^A. \tag{7}$$

IV. OPTIMIZATION

Since we are working with a routing problem, our main goal is to find the optimal allocation of flow for our network given an objective function to minimize. Our decision variables for the optimization are the flows of trucks f_p^T along the given paths $p \in \mathcal{P}^R$. As mentioned before, the flows of drones f_p^D along the aerial paths $p \in \mathcal{P}^A$ can be seen as linear functions of our decision variable and hence do not have to be explicitly controlled in the optimization. We will now explain the objective function used, as well as the constraints required by our model.

A. Objective Function

We want to find the optimal routing for a parcel delivery system that minimizes both parcel latency and societal costs. Because of this, our objective function consists of two components, the average latency experienced by parcels L, and the societal cost incurred on the system by the parcel delivery service L^S. We express our objective function as:

$$J(f^T) = \gamma L + (1 - \gamma)L^S, \tag{8}$$

where we use $0 < \gamma < 1$ as a weight that determines the relative importance of the parcel latency and societal cost objectives.

B. Constraints

The first constraint of our model is the operational cost defined in (7). This is seen as an inequality constraint: $C(f^T) \leq C_0$, where C_0 is a constant denoting the maximal operational cost allowed in dollars per hour. Additionally, we need to fulfill the system's parcel demand while maintaining conservation of flow within the network by satisfying (1) and (2).

C. Overall Optimization Problem

Our overall optimization problem can be written as:

$$\min_{f^T} \; J(f^T) = \gamma L + (1 - \gamma)L^S$$
$$\text{subject to} \;\; f_p^T \geq 0 \qquad \forall p \in \mathcal{P}^R \,,$$
$$C(f^T) \leq C_0 \,, \tag{9}$$
$$d_v^R(f^T) = m\Big(\sum_{e \in \mathcal{U}_v^R} f_e^T - \sum_{e \in \mathcal{D}_v^R} f_e^T \Big) \,,$$
$$\text{and} \;\; d_v^A(f^T) = d_v - d_v^R(f^T), \qquad \forall v \in \mathcal{N} \,.$$

We will show in Section VI how the above formulation in (9) can be expressed as a quadratic convex objective function with linear equality and inequality constraints.

V. SUMO SIMULATIONS

Using the open source traffic simulation package "SUMO" [12], we derive expressions for latency as a function of the stopping and non-stopping flow components on a given link. In this section, we will briefly describe the simulation setup used and subsequently show the analyzed results.

A. Setup

For all of our simulations, we used two vehicle types - cars that do not stop at delivery locations, and trucks that stop at delivery locations. The vehicle parameters used throughout the simulations are listed in Table I.

TABLE I: Vehicle Parameters

Vehicle	Car	Truck
Length m [33]	4.00	7.82
Max. Speed $\frac{km}{h}$ [34]	160.93	96.56
Max. Acceleration $\frac{m}{s^2}$ [34]	2.87	1.00
Max. Deceleration $\frac{m}{s^2}$ [34]	4.33	0.88

With these vehicle types we simulated two variants of roads, one with two lanes and the other with three lanes, placing equally spaced truck stops along both roads. For each road, we simulated over a set of two parameters - the ratio of stopping to non stopping vehicles and the total flow of vehicles. We varied the ratio from 0.001 to 0.1 since that is the regime of practical interest, and for each value we increased the total flow in increments of 10 cars per hour until reaching a prescribed maximum value. The road parameters used throughout the simulations are listed in Table II. After collecting the data for the following simulations, we analyzed our results to see how latency was affected by the input parameters.

TABLE II: Road Parameters

Road	Two Lane	Three Lane
Length m	500	2000
Speed Limit $\frac{km}{h}$	48.28	80.47
Number of Stops	25	100
Duration $minutes$	60	90
Max. Flow $\frac{vehicles}{hour}$	3000	5000

B. Analysis

For each simulation, we looked at the travel times experienced by all vehicles, averaging this value to compute the
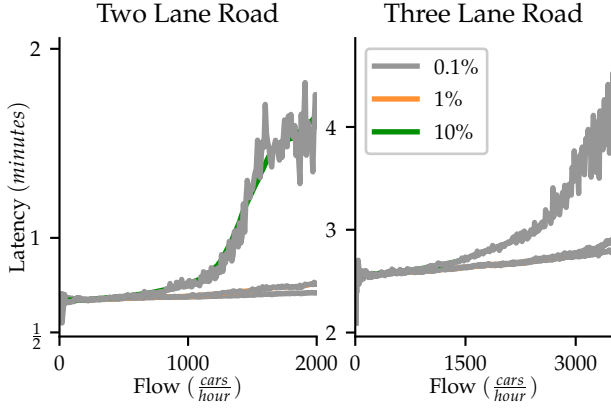
Fig. 4: The three lines in each plot represent simulation results for different ratios of truck flow to total flow. As the total flow of vehicles increases, so does the latency, with larger ratios contributing to sharper increases. From the two subplots we can see that two lane roads are proportionally more affected by stopping trucks as compared to three lane roads.

road latency given the current ratio and total flow. From Fig. 4, we can see that latency is much larger when there is a bigger portion of stopping trucks on the road, and that this effect is more severe for two lane roads. We also see that for smaller ratio values, the relationship between latency and truck flow can be modeled as linear, since in this regime each truck creates a constant amount of delay. Since in practice, we expect the flow of stopping trucks f_e^T along a given edge e to be much smaller than the nominal vehicular flow f_e^0 along that edge, we consider this operating regime and approximate our edge latencies ℓ_e^R as linear with respect to the two components of flow:

$$\ell_e^R = \langle \omega_e, [1, f_e^R, f_e^R + f_e^0] \rangle, \qquad (10)$$

where ω_e corresponds to the 3-dimensional weight vector associated with edge $e \in \mathcal{E}^R$ and $\langle a, b \rangle$ is used to denote the inner product between vectors a and b. As shown in

Fig. 3, this is equivalent to a 3-dimensional linear plane for each simulated road. It is important to note that our weight vectors have strictly positive terms, since latency is monotonically increasing with respect to the flow components. Using our SUMO simulations, we are able to define affine transformations that map from our network flow components to latencies.

VI. SIMULATION STUDIES

Now that we have derived the functions used to compute our edge latencies ℓ_e^R, we can go back to solving our overall optimization problem shown in (9). Using our expressions for average latency in (3), (4), (5), (6), and our derived latency model in (10), we can express our objective function in standard quadratic form:

$$J(f^T, \gamma) = (f^T)^\intercal Q f^T + a^\intercal f^T, \qquad (11)$$

where we use $\intercal$ to denote the transpose operation, and γ has been absorbed into the expressions for matrix Q and vector a.

For a given problem setup, Q is a diagonal matrix where the i'th row's diagonal term is γ times the sum of the second and third components of ω_e for all $e \in p$, where p is the corresponding path for row i. Since all components of ω_e are positive, Q is positive-definite. Vector a's construction requires many operations, and hence is left out for readability purposes as it can be derived using the equations provided. We go on to show a small example that visualizes how we construct our network model and optimize. Following this, we preform a case study on the Sioux Falls transportation network, shown in 7, to demonstrate the scalability of our methods.

A. Complexity and Scalability

Since the problem posed in (9) is defined by a convex quadratic objective function, it can be efficiently solved using quadratic programming (QP) techniques. Inspired by
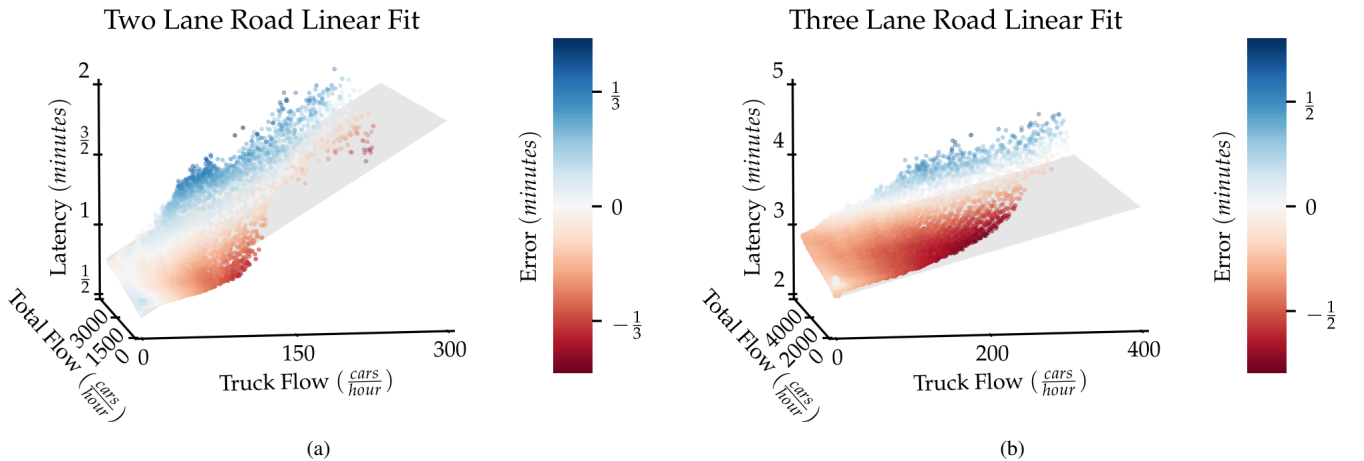

(a)

(b)

Fig. 3: The two plots (a) and (b) depict the best fit linear planes derived for our two-lane road and three-lane road respectively. The z-dimension represents latency, while the xy-dimensions correspond to our two flow components. Colors are used to visualize the error between the simulation results and the best fit planes. We can see that our affine transformations slightly over-estimate the latency when total flow is small, and under-estimate the latency when total flow is large. We emphasize that, because trucks account for a small portion of the total flow, we operate in the regime where errors are small.

Karmarkar's Linear Programming algorithm [35], multiple interior-point algorithms have adapted his method for the QPs, showing convergence to the optimal solution in $O(n^3L)$ arithmetic operations [36], [37], [38], where n is the dimension of the decision variable and L is the number of binary symbols needed to record the input information.

In our specific case, n is the number of paths considered in the problem $|\mathcal{P}^R|$, and L is the number of bits required to store the information in Q, a from (11) as well as the constraints defined in (9). Since the number of paths between two nodes exhibits factorial scaling with the number of links between them, the total number of paths can scale worse than exponentially. Fortunately, we can restrict ourselves to consider paths up to a certain length, largely limiting the computational overhead.

B. Example Network

For our example network we choose a set of four nodes $\mathcal{N}$ and two sets of edges, $\mathcal{E}^R$ and $\mathcal{E}^A$, building our road and aerial networks, G^R and G^A, as shown in Fig. 6. The nominal flows f_e^0 and parcel demands d_v are shown in Fig. 6, where $\beta = 14,000$ vehicles per hour is used to compute the societal latency. We use $c^T = 30$ dollars per hour and $c^D = 0.5$ dollars per hour as our operational costs for trucks and drones respectively, allowing each truck to carry a total of $m = 125$ parcels [28]. Finally, we set the drone speed as 25 kilometers per hour, as this should allow for the longest flight distance [32].

With this setup, we optimize for our objective function using multiple values of γ to shift priority between minimizing societal latency and minimizing parcel latency. For this example, we chose to use the full set of available paths
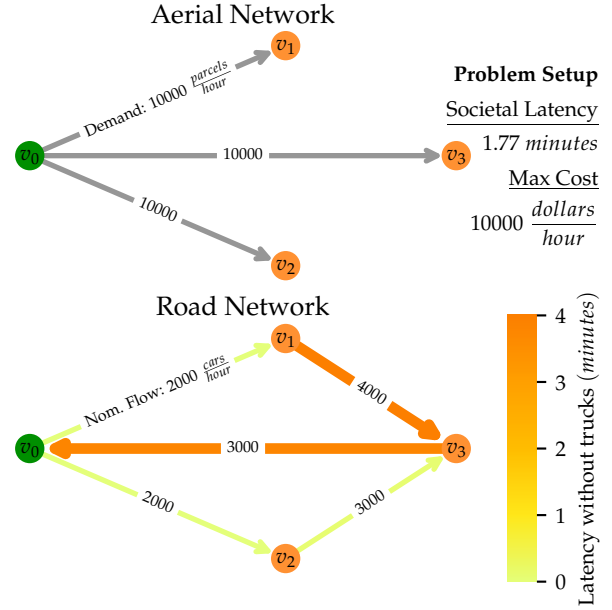

Fig. 6: This figure depicts the setup for our small network example, note that it is not drawn to scale. The top graph represents G^A, labeling the parcel demands for each node along the corresponding edge. The bottom graph represents G^R, labeling the nominal flows of vehicles along the edges. The thicker edges represent three lane roads that are 2 kilometers long, while the thinner edges represent two lane roads that are 0.5 kilometers long.

as our decision variable, depicting our results in Table III.

As expected, we can see that utilizing drones helps with minimizing both societal latency and parcel latency. More specifically, we see that with drones we can prioritize minimizing for societal latency by decreasing γ, meanwhile
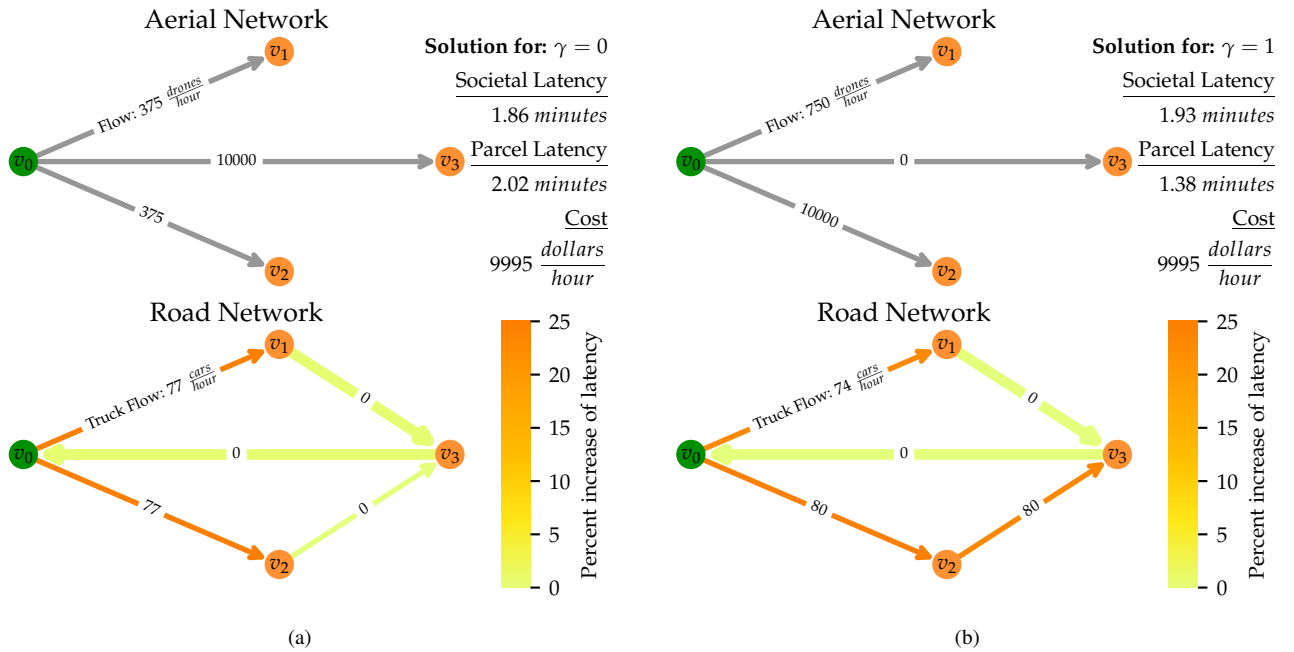

Fig. 5: The two figures (a) and (b) show the solutions to our optimization problem for different values of γ. Societal and parcel latency are calculated according to (6) and (3) respectively. The edge labels in the aerial networks are the flows of drones along the paths. The edge labels in the road networks are the flows of trucks along the edges, which are calculated post hoc using the solution's path flows.

TABLE III: Example Network Results

Setup	Without Drones	With Drones
$\gamma = 0$	1.98, 2.29	1.86, 2.02
$\gamma = \frac{1}{2}$	1.98, 1.81	1.89, 1.70
$\gamma = 1$	1.98, 1.63	1.93, 1.38

TABLE III: This table shows the results for our example network using multiple values of γ. Each entry includes two values: societal latency and parcel latency, both of which are reported in minutes.

without drones, decreasing γ has little effect on societal latency. This is due to the chosen setup of our example, showcasing that trucks will in general have a negative impact on congestion within the road network.

The solutions for $\gamma = 0$ and $\gamma = 1$ are depicted in Fig. 5, where the edge labels correspond to the drone and truck flows derived from our solution. We can see that setting γ to 0 routes the drones to v_3, the farthest node. This causes an increase in parcel latency, but keeps trucks off of the road connecting v_2 to v_3. On the other hand, setting γ to 1 routes the drones to v_2 and lets the trucks make faster deliveries to v_3 instead. This reduces the overall parcel latency, but increases societal latency due to the added congestion on the road connecting v_2 to v_3.

C. Sioux Falls Case Study

For our case study, we use the Sioux Fall transportation network shown in Fig. 7 since it has been widely used as a standard test for traffic assignment [39]. The network contains 24 nodes and 76 directed edges, where we choose node v_{10} as our central delivery hub. Using the reported values for the best known edge flows and the edge capacities, we scale our nominal flows f_e^0 to be proportional to the ratios between edge flows and edge capacities. We choose 24 of the 76 edges to be three lane roads since their capacities are over $10,000$ vehicles per hour, and scale the weights ω_e for all edges $e \in \mathcal{E}^R$ based on the physical distances between the nodes each edge connects. We consider all 1133 paths that consist of eight or less edges, and optimize for our objective function using parcel demands of 5000 parcels per hour for each node, setting the total nominal flow β to $81,000$ vehicles per hour, and allowing a maximum operational cost of $50,000$ dollars per hour. We summarize the results of our case study in Table IV.

First of all, we see a similar trend as depicted in our example network: utilizing drones minimizes both societal latency and parcel latency. In addition, we see that drones have a greater affect on parcel latency than was depicted by our example network. Since our case study's road network consists of an actual grid rather than direct links between nodes, we see the advantages drones have due to their ability to utilize direct paths between nodes. Overall, our framework quantifies the potential drones have on mitigating road congestion and solves for the path routing needed to minimize a chosen combination of societal latency and parcel latency, while remaining computationally feasible to scale due to its reliance on convex quadratic optimization.
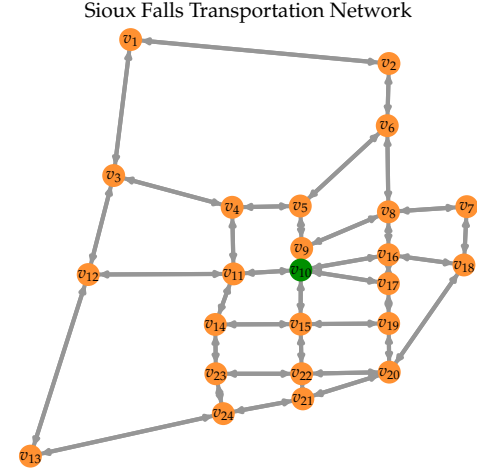
Sioux Falls Transportation Network

Fig. 7: The Sioux Falls transportation network is depicted above. The green node v_{10} represents our delivery hub's location.

TABLE IV: Case Study Results

Setup	Without Drones	With Drones
$\gamma = 0$	9.33, 10.97	8.70, 7.79
$\gamma = \frac{1}{2}$	9.39, 8.94	8.82, 7.33
$\gamma = 1$	9.51, 8.29	9.00, 7.20

TABLE IV: This table shows the results for our case study using multiple values of γ. Each entry includes two values: societal latency and parcel latency, both of which are reported in minutes.

VII. CONCLUSION

We proposed a network model for a congestion-aware Bi-modal delivery system incorporating drones. Using our simulation results, we constructed a convex optimization framework which can solve for the optimal routing of our delivery system. Finally, we showed the scalability of our framework by applying it on the well studied Sioux Falls transportation network. Below we discuss the limitations of our work, including directions for possible improvement.

An inaccuracy of our framework is the planar fit latency models depicted in Fig. 3. One way to address this issue is by utilizing a piece-wise linear fit instead. This addition would make our main objective function piece-wise quadratic, meaning we can no longer rely on simple quadratic programming solvers for our optimization. We point interested readers to a study of such piece-wise quadratic solvers, leaving this extension to future work [40].

In addition to the inaccuracies of our latency estimates, one may view the simplicity of our proposed networks as a limitation as well. We discourage this perspective, since the inherent simplicity is what allows for the construction of a scalable and convex optimization framework. As pointed out before, prior works have shown lack of applicability in large scale networks due to their reliance on heuristic techniques and MILP formulations to solve for routing policies [17], [18], [19]. On the other hand, our proposed network model and optimization framework lets us scale to arbitrarily large networks, allowing delivery systems to find routing policies for their trucks and drones, and evaluate the effects different routing policies can have on both their parcel delivery latency and the traffic conditions of the roads.

REFERENCES

[1] J. Holguín-Veras, J. Amaya, I. Sánchez-Díaz, M. Browne, and J. Wojtowicz, "State of the art and practice of urban freight management part ii: Financial approaches, logistics, and demand management," *Transportation Research Part A Policy and Practice*, vol. 137, pp. 383–410, 07 2020.

[2] F. Ali, "Us ecommerce grows 44.0 in 2020," Jan 2021. [Online]. Available: https://www.digitalcommerce360.com/article/us-ecommerce-sales/

[3] D. Kong, X. Guo, B. Yang, and D. Wu, "Analyzing the impact of trucks on traffic flow based on an improved cellular automaton model," *Discrete Dynamics in Nature and Society*, vol. 2016, 01 2016.

[4] M. Al Eisaeia, S. Moridpourb, and R. Tay, "Heavy vehicle management: Restriction strategies," *Transportation Research Procedia*, vol. 21, pp. 18–28, 2017, international Symposia of Transport Simulation (ISTS) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146517302119

[5] R. G. Thompson, *Vehicle Oriented Initiatives for Improving the Environmental Performance of Urban Freight Systems*. Cham: Springer International Publishing, 2015, pp. 119–129. [Online]. Available: https://doi.org/10.1007/978-3-319-17181-4_7

[6] L. Dablanc and J.-P. Rodrigue, "The geography of urban freight. in : The geography of urban transportation," 01 2014.

[7] S. Choudhury, K. Solovey, M. J. Kochenderfer, and M. Pavone, "Efficient large-scale multi-drone delivery using transit networks," 2021.

[8] N. Kafle, B. Zou, and J. Lin, "Design and modeling of a crowdsource-enabled system for urban parcel relay and delivery," *Transportation Research Part B: Methodological*, vol. 99, pp. 62–82, 05 2017.

[9] M. Joerss, F. Neuhaus, and J. Schröder, "How customer demands are reshaping last-mile delivery," Nov 2020. [Online]. Available: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/how-customer-demands-are-reshaping-last-mile-delivery

[10] SESARJU, "Smart atm u-space," Jan 2021. [Online]. Available: https://www.sesarju.eu/U-space

[11] DPDHL, "Dhl parcelcopter." [Online]. Available: https://www.dpdhl.com/en/media-relations/specials/dhl-parcelcopter.html

[12] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE, 2018. [Online]. Available: https://elib.dlr.de/124092/

[13] R. Merkert and J. Bushell, "Managing the drone revolution: A systematic literature review into the current use of airborne drones and future strategic directions for their effective control," *Journal of air transport management*, vol. 89, p. 101929, 10 2020.

[14] J. Lee, "Optimization of a modular drone delivery system," 04 2017, pp. 1–8.

[15] I. Hong, M. Kuby, and A. Murray, "A deviation flow refueling location model for continuous space: A commercial drone delivery system for urban areas," in *Advances in Geocomputation*, D. A. Griffith, Y. Chun, and D. J. Dean, Eds. Cham: Springer International Publishing, 2017, pp. 125–132.

[16] M. Shavarani, M. Ghadiri Nejad, F. Rismanchian, and G. Izbirak, "Application of hierarchical facility location problem for optimization of a drone delivery system: a case study of amazon prime air in the city of san francisco," *The International Journal of Advanced Manufacturing Technology*, vol. 95, pp. 3141–3153, 04 2018.

[17] J. Caceres-Cruz, P. Arias, D. Guimarans, D. Riera, and A. A. Juan, "Rich vehicle routing problem: Survey," *ACM Comput. Surv.*, vol. 47, no. 2, Dec. 2014. [Online]. Available: https://doi.org/10.1145/2666003

[18] A. Otto, N. Agatz, J. Campbell, B. Golden, and E. Pesch, "Optimization approaches for civil applications of unmanned aerial vehicles (uavs) or aerial drones: A survey," *Networks*, vol. 72, no. 4, pp. 411–458, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/net.21818

[19] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, p. 70–85, Jan 2017. [Online]. Available: http://dx.doi.org/10.1109/TSMC.2016.2582745

[20] A. Karak and K. Abdelghany, "The hybrid vehicle-drone routing problem for pick-up and delivery services," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 427–449, 2019.

[21] R. Liu and Z. Jiang, "A hybrid large-neighborhood search algorithm for the cumulative capacitated vehicle routing problem with time-window constraints," *Applied Soft Computing*, vol. 80, pp. 18–30, 2019.

[22] T. Vidal, G. Laporte, and P. Matl, "A concise guide to existing and emerging vehicle routing problem variants," *European Journal of Operational Research*, vol. 286, no. 2, pp. 401–416, 2020.

[23] H. Bast, D. Delling, A. Goldberg, M. Müller-Hannemann, T. Pajor, P. Sanders, D. Wagner, and R. F. Werneck, "Route planning in transportation networks," 2015.

[24] F. M. Bergmann, S. M. Wagner, and M. Winkenbach, "Integrating first-mile pickup and last-mile delivery on shared vehicle routes for efficient urban e-commerce distribution," *Transportation Research Part B: Methodological*, vol. 131, pp. 26–62, 2020.

[25] S. Goswami and P. Tallapragada, "One-shot coordination of feeder vehicles for multi-modal transportation," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 1714–1719.

[26] V. Reis, J. F. Meier, G. Pace, and R. Palacin, "Rail and multi-modal transport," *Research in Transportation Economics*, vol. 41, no. 1, pp. 17–30, 2013.

[27] S. Choudhury, J. P. Knickerbocker, and M. J. Kochenderfer, "Dynamic real-time multimodal routing with hierarchical hybrid planning," 2019.

[28] M. Kim and E. T. Matson, "A cost-optimization model in multi-agent system routing for drone delivery," in *Highlights of Practical Applications of Cyber-Physical Multi-Agent Systems*, J. Bajo, Z. Vale, K. Hallenborg, A. P. Rocha, P. Mathieu, P. Pawlewski, E. Del Val, P. Novais, F. Lopes, N. D. Duque Méndez, V. Julián, and J. Holmgren, Eds. Cham: Springer International Publishing, 2017, pp. 40–51.

[29] A. M. Ham, "Integrated scheduling of m-truck, m-drone, and m-depot constrained by time-window, drop-pickup, and m-visit using constraint programming," *Transportation Research Part C: Emerging Technologies*, vol. 91, pp. 1–14, 2018.

[30] S. Ferrandez, T. Harbison, T. Weber, R. Sturges, and R. Rich, "Optimization of a truck-drone in tandem delivery network using k-means and genetic algorithm," *Journal of Industrial Engineering and Management*, vol. 9, p. 374, 04 2016.

[31] C. C. Murray and A. G. Chu, "The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 86–109, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X15000844

[32] DJI, "Dji mavic 2 specifications sheet." [Online]. Available: https://www.dji.com/mavic-2/info#specs

[33] B. Sharpe and F. Rodriguez, "Market analysis of heavy-duty commercial trailers in europe," 09 2018.

[34] P. Bokare and A. Maurya, "Acceleration-deceleration behaviour of various vehicle types," *Transportation Research Procedia*, vol. 25, pp. 4733–4749, 2017, world Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146517307937

[35] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica (Print)*, vol. 4, no. 4, pp. 373–395, 1984.

[36] D. Goldfarb and S. Liu, "An o (n 3 l) primal interior point algorithm for convex quadratic programming," *Mathematical programming*, vol. 49, no. 1, pp. 325–340, 1990.

[37] Y. Ye and E. Tse, "An extension of karmarkar's projective algorithm for convex quadratic programming," *Mathematical programming*, vol. 44, no. 1, pp. 157–179, 1989.

[38] R. D. Monteiro and I. Adler, "Interior path following primal-dual algorithms. part i: Linear programming," *Mathematical programming*, vol. 44, no. 1, pp. 27–41, 1989.

[39] B. Stabler, "Transportation networks for research core team." [Online]. Available: https://github.com/bstabler/TransportationNetworks

[40] Y. Cui, T.-H. Chang, M. Hong, and J.-S. Pang, "A study of piecewise linear-quadratic programs," 2018.